Comparative genomic analysis of two strains of human adenovirus type 3 isolated from children with acute respiratory infection in southern China

Qiwei Zhang,[1,2] Xiaobo Su,[3] Sitang Gong,[2] Qiyi Zeng,[2] Bing Zhu,[2] Zaohe Wu,[3] Tao Peng,[4] Chuyu Zhang[1] and Rong Zhou[2,3]

Correspondence
Chuyu Zhang
Zhang_whu@yahoo.com.cn

[1]State Key Laboratory of Virology, College of Life Sciences, Wuhan University, Wuhan 430072, China

[2]Central Laboratory, Guangzhou Children's Hospital, Guangzhou 510120, China

[3]South China Sea Institute of Oceanology, LED, Chinese Academy of Sciences, Guangzhou 510301, China

[4]Guangzhou Institute of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510663, China

Human adenovirus type 3 (HAdV-3) is a causative agent of acute respiratory disease, which is prevalent throughout the world, especially in Asia. Here, the complete genome sequences of two field strains of HAdV-3 (strains GZ1 and GZ2) isolated from children with acute respiratory infection in southern China are reported (GenBank accession nos DQ099432 and DQ105654, respectively). The genomes were 35 273 bp (GZ1) and 35 269 bp (GZ2) and both had a G + C content of 51 mol%. They shared 99 % nucleotide identity and the four early and five late regions that are characteristic of human adenoviruses. Thirty-nine protein- and two RNA-coding sequences were identified in the genome sequences of both strains. Protein pX had a predicted molecular mass of 8·3 kDa in strain GZ1; this was lower (7·6 kDa) in strain GZ2. Both strains contained 10 short inverted repeats, in addition to their inverted terminal repeats (111 bp). Comparative whole-genome analysis revealed 93 mismatches and four insertions/deletions between the two strains. Strain GZ1 infection produced a typical cytopathic effect, whereas strain GZ2 did not; non-synonymous substitutions in proteins of GZ2 may be responsible for this difference.

INTRODUCTION

Adenoviruses (AdVs) are responsible for 5–10 % of lower respiratory tract infections in infants and children and infect a very broad spectrum of hosts, including cattle, duck, possum, dog, tree shrew, fish, frog, corn snake, and equine, ovine, porcine and simian animals (Davison *et al.*, 2000; Farkas *et al.*, 2002; Kovács *et al.*, 2003). They can be divided into four genera, *Atadenovirus*, *Aviadenovirus*, *Mastadenovirus*, *Siadenovirus*, and unassigned species (Benkő *et al.*, 2000, 2002; Shenk, 2001; Davison *et al.*, 2003; Kovács *et al.*, 2003). Since the first human adenovirus (HAdV) was isolated (Rowe *et al.*, 1953), 51 different HAdV serotypes have been identified within the genus *Mastadenovirus* and they can be classified into six species (HAdV-A to -F) based on a variety of parameters, including oncogenicity in rodents, electrophoretic mobility (Wadell, 1979) and DNA

or genome identity (Garon *et al.*, 1973; Green *et al.*, 1979; Wadell *et al.*, 1980; Wadell, 1984; De Jong *et al.*, 1999), as well as the classical gold standards of serum neutralization and haemagglutination-inhibition tests (Davison *et al.*, 2003).

HAdV-B species have been divided further into two groups: B1, including HAdV-3, -7, -16, -21 and -50, and SAdV-21; and B2, including HAdV-11, -14, -34 and -35 (Wold *et al.*, 1979; Stone *et al.*, 2003). HAdV-B group B1 viruses have been isolated from patients with febrile respiratory disease, especially fatal acute respiratory disease (ARD) (Hierholzer, 1995; Erdman *et al.*, 2002). Members of group B2, with the exception of HAdV-11a and -14 (Van der Veen, 1963; Mei *et al.*, 1998), are associated with persistent infections of kidney and urinary tract (Myerowitz *et al.*, 1975; Shields *et al.*, 1985).

Of the 51 HAdV serotypes, about one-third are associated with human diseases. Adenovirus infections can occur endemically or as outbreaks. HAdV-B group B1 (HAdV-3,

HAdV-7 and, less frequently, HAdV-21) and HAdV-4 have been the causative agents in epidemic outbreaks of respiratory disease in Europe, America, Oceania and Asia (Herbert *et al.*, 1977; Martone *et al.*, 1980; Lewis *et al.*, 2004; Frantzidou *et al.*, 2005). Viruses of group B1 (HAdV-3, -7 and -21) can occasionally infect tissues of the central nervous system and cause aseptic meningitis, meningoencephalitis and encephalitis (Chany *et al.*, 1958; Faulkner & Van Rooyen, 1962; Similä *et al.*, 1970).

Epidemic outbreaks of ARD caused by HAdV-4 and HAdV-7 among American and Canadian basic military trainees have been controlled by the introduction of effective live enteric-coated oral vaccines since the 1970s (Dudding *et al.*, 1972); however, the manufacture of HAdV vaccines was discontinued in 1996. HAdV-3, first isolated by investigators in the Walter Reed Army Institute of Research from patients with ARD at Fort Leonard Wood (MO, USA) in the winter of 1952–1953 (http://history.amedd.army.mil/booksdocs/historiesofcomsn/section1.htm), is widely prevalent all over the world (Herbert *et al.*, 1977; Martone *et al.*, 1980; Ryan *et al.*, 2002; Frantzidou *et al.*, 2005), especially in Asia (Itakura *et al.*, 1990; Itoh *et al.*, 1999; Hong *et al.*, 2001; Kim *et al.*, 2003; Li *et al.*, 2004). However, no efficient vaccine against HAdV-3 has been developed. HAdV infections are highly contagious and common in dense and close populations, such as military training venues and day-care centres. The population of Asia is large and often dense, especially in China and Japan. Consequently, epidemics of ARD caused by HAdVs occur at high frequency. In July 2004, more than 200 children from an infant school were infected with HAdV-3 in Guangzhou, southern China (Zhu *et al.*, 2005).

In Asia, multiple HAdV-3 genome types have been identified by restriction-enzyme analysis (Itakura *et al.*, 1990; Itoh *et al.*, 1999; Kim *et al.*, 2003). In China, the dominant genome type from 1962 to 1988 was HAdV-3a2, with occasional isolates of HAdV-3a4, HAdV-3a5 and HAdV-3a6 (Li & Wadell, 1988; Li *et al.*, 1996). In Japan, the dominant genome type from 1983 to 1991 was HAdV-3a, with occasional isolates of HAdV-3a8 and HAdV-3c (Itakura *et al.*, 1990; Mizuta *et al.*, 1994; Shiao *et al.*, 1996). In Seoul, Korea, six new variants, HAdV-3a13 to HAdV-3a18, were found from 1990 to 2000 (Kim *et al.*, 2003). Genome types may vary by location and time of isolation; some genome types may be associated with greater virulence (Kajon *et al.*, 1990, 1996). It is therefore important to understand the genomics and bioinformatics of human disease-relevant HAdVs of group B1.

Since the first HAdV genome sequence was reported (HAdV-2) (Roberts *et al.*, 1984, 1986), the complete genome sequences of 21 members of the genus *Mastadenovirus* have been released, with at least one from each species. For HAdV-B, to date, genomes of HAdV-7, -11, -21, -35 and -50 have been deposited in GenBank/EMBL (Gao *et al.*, 2003; Mei *et al.*, 2003; Stone *et al.*, 2003; Vogels *et al.*, 2003; Roy *et al.*, 2004; Purkayastha *et al.*, 2005). The complete genomic sequence of HAdV-3 has not been reported previously. In this report, two complete and annotated genome

sequences of HAdV-3a, strains Guangzhou01 (GZ1) and Guangzhou02 (GZ2), are described (GenBank accession nos DQ099432 and DQ105654, respectively); open reading frames and non-coding motifs were also analysed and compared.

Strain GZ1 infection produced a typical cytopathic effect (CPE), whereas strain GZ2 did not. The genome organization of both strains is similar to that observed in other members of HAdV-B. Bioinformatics provides an insight into the biology of HAdV-3 and raised our interest in the CPE difference caused by the two strains. The clinical application of HAdV-2- and HAdV-5-based gene-transfer vectors has been hampered because of pre-existing immunity against HAdV-2 and HAdV-5, which could affect the efficacy and even safety of adenovirus vector administration. HAdV-3, unlike HAdVs that do not belong to HAdV-B, has no proven affinity for the coxsackievirus–adenovirus receptor (CAR) (Roelvink *et al.*, 1998). This receptor diversity implies that HAdV-3 has a different tropism from CAR-interacting AdVs and could provide an alternative to HAdV-5-based gene-transfer vectors (Havenga *et al.*, 2002; Sirena *et al.*, 2004).

METHODS

Cells and virus strains. HAdV-3 strains GZ1 and GZ2 were isolated from nasal aspirates of children with clinical evidence of ARD in January 2005 and July 2004, respectively. The child from which strain GZ2 originated had the symptoms of pharyngeal conjunctivitis and the other child had fever and bronchitis. Nasal aspirate specimens were inoculated into HEp-2, MDCK and HeLa culture tubes with an atmosphere of 5 % (v/v) carbon dioxide in Dulbecco's minimum essential medium supplemented with 100 IU penicillin ml^{-1}, 100 µg streptomycin ml^{-1} and 2 % (v/v) fetal calf serum. The culture tubes were observed for 3–4 weeks for CPE and identified by a neutralization assay with type-specific reference antisera raised in rabbits by conventional procedures (Hierholzer, 1995). Type-specific primers designed to the hypervariable regions (HVRs) of the HAdV hexon were also utilized to correctly identify HAdV-3, -4, -7 and -11. The following primers were used: HAdV-3-specific, HAdV-3S (5′-AAGACATTACCACTACTGAAGGAGAAG-3′) and HAdV-3A (5′-CGCTAAAGCTCCTGCAACAGCA-3′); HAdV-4-specific, HAdV-4S (5′-GGTAGCTGCCATGCCAGGTG-3′) and HAdV-4R (5′-CATA-GTTAGGAGTGGCGGCGG-3′); HAdV-7-specific, HAdV-7S (5′-GGGAAAGACATTACTGCAGACAAC-3′) and HAdV-7R (5′-GGC-GAAAAAGCGTCAGCAG-3′); and HAdV-11-specific, HAdV-11S (5′-AGGAACACGTAACAGAAGAGGAAACC-3′) and HAdV-11R (5′-TAGCTTCGGAACTTGTGTCTTCTGTT-3′).

Preparation of viral DNA and genome-type analysis. Virus was propagated in HEp-2 cells and viral DNA was extracted by using a previously described method (Shinagawa *et al.*, 1983). Purified HAdV genome DNA was digested by restriction enzymes (*Bam*HI, *Eco*RI, *Hin*dIII, *Sal*I and *Sma*I; TaKaRa). HAdV strains were genome-typed by comparing the restriction profiles with those of prototype and other genome types described in the literature and according to the genome-type denomination system (Li & Wadell, 1988; Golovina *et al.*, 1991; Li *et al.*, 1996; Kim *et al.*, 2003).

DNA cloning and sequencing. The restriction fragments of HAdV genome DNA digested with *Hin*dIII, *Eco*RI, *Bam*HI or *Sma*I were purified with QIAquick Gel Extraction kits (Qiagen) and cloned into pBlueScript SK(+) vectors. The entire HAdV-3 genome DNA was resequenced using primers from initially sequenced

regions cloned in vectors and various HAdV-3 gene sequences archived in GenBank/EMBL. Template DNA (0·1–1·0 µg per reaction) was further purified by passing through Mini Spin Columns (Qiagen). The sequencing reaction was carried out by using an ABI Prism BigDye Terminator v3.1 Cycle Sequencing Ready Reaction kit with AmpliTaq DNA polymerase on an ABI 3730 DNA sequencer (Applied Biosystems). All of the reported sequences are the result of at least three sequencing reactions.

Direct sequencing of inverted terminal repeat (ITR) ends. The 5′ and 3′ ends of the linear HAdV-3 genome were sequenced directly on an ABI 3730 DNA sequencer (Applied Biosystems) with the repurified genomic DNA as templates. Primers were designed from newly obtained internal sequences.

Genome annotation and sequence analysis. The sequences were assembled with SEQMAN software from the Lasergene package (DNAStar) and SEQUENCHER 4.1.4 (Gene Codes). Genome annotation provided an additional layer of sequence quality control. Unresolved and ambiguous sequences were resequenced with primers close to the regions in question.

General features of the HAdV-3 genome sequences were revealed by using the University of Wisconsin Genetics Computer Group (GCG) package (SEQWEB v. 2). The genome sequence was annotated with the annotation protocol used for HAdV-1 genome analysis (Lauer et al., 2004) by first dividing the sequence into contiguous 1 kb non-overlapping segments. Briefly, these segments were queried systematically against the non-redundant NCBI database using the program BLASTX of the BLAST suite of sequence-alignment software (Altschul et al., 1990). Default parameters of word size = 3 and expectation = 10, with the BLOSUM62 substitution matrix and with gap penalties of 11 (existence) and 1 (extension), were applied to these analyses. Low-complexity sequences were filtered out of the queries, as per the BLAST algorithm.

GENSCAN 1.0 and GENOMESCAN were used for theoretical gene predictions (Yeh et al., 2001). They were useful for identifying exons from the coding sequences where exon–intron borders were difficult to determine. Other splice site-finder programs [WISE2 (http://www.ebi.ac.uk/Wise2/advanced.html) and SPLICEPREDICTOR (Brendel & Kleffe, 1998)] were used to find splice-donor and -acceptor sites with the highest score. In parallel, novel sequences or 'hypothetical proteins'

were also identified by using FGENESV, software for predicting potential genes in viral genomes (http://www.softberry.com/berry.phtml?topic = index&group = programs&subgroup = gfindv) and GENEMARK v. 2.4, a Hidden Markov method-based gene-prediction software (Besemer & Borodovsky, 1999). In these annotations, although FGENESV had a slightly higher accuracy than the others, none of them were completely comprehensive or accurate in predicting putative genes. To visualize the annotation progress, the genome-annotation and -editor tool ARTEMIS was used to expedite genome annotation (Berriman & Rutherford, 2003).

Whole-genome alignment and comparisons of the sequences from HAdVs were performed by using the dot-plot software Advanced PipMaker (http://pipmaker.bx.psu.edu/cgi-bin/pipmaker?advanced), which aligns long genomic DNA sequences quickly and with good sensitivity (Schwartz et al., 2000).

RESULTS AND DISCUSSION

Confirmation of serotype and genome type

Typical CPE was found in all cells inoculated with strain GZ1, but not with strain GZ2. Both virus strains were neutralized specifically by HAdV-3-neutralizing rabbit immune serum. Antisera against the closely related HAdV-7 cross-reacted only slightly. PCR assay also indicated that both HAdV strains were serotype 3, for only PCR with primers specific to HAdV-3 could obtain a product of 314 bp. Further genome-typing results of restriction profiles made it clear that both strains were genome type HAdV-3a.

General characteristics of the HAdV-3 genome sequence

The genome sequences of HAdV-3 strains GZ1 and GZ2 were annotated to identify biological features. This was facilitated by using reference genomes from the recently determined HAdV-11 (GenBank/EMBL accession nos AF532578 and AY163756) and HAdV-7 (GenBank/EMBL
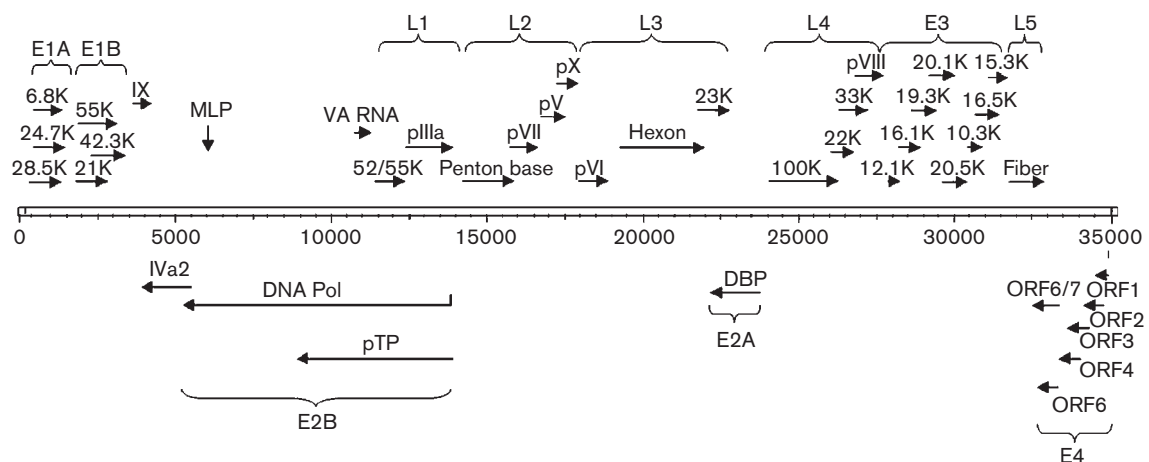
Fig. 1. Genomic organization and transcription map of HAdV-3 strains GZ1 and GZ2. Arrows indicate the locations of coding regions. Early and late transcription units are shown with brackets. Abbreviations: DBP, DNA-binding protein; pTP, terminal protein precursor; MLP, major late promoter.

Table 1. HAdV-3 strains GZ1 and GZ2 genome-sequence annotation and comparison

Non-coding motifs and coding regions are identified for HAdV-3 strains GZ1 and GZ2. Their proteins and putative functions are indicated. The nucleotide positions of the start and stop codons and the applicable splice sites are noted (5′→3′ direction). Functionality embedded within the complementary strand and coding sequences transcribed from the complementary strand are indicated by 'c', e.g. (3925–3930)c. An asterisk indicates that the same position is found in strain GZ2 as in strain GZ1. − indicates that the strain does not include the corresponding protein.

Region	Product	Genome location	
		Strain GZ1	Strain GZ2
	ITR	1–111	*
	DNA Pol–pTP binding site	9–18	*
E1A	TATA box for E1A	480–485	*
E1A	6·8 kDa protein	576–647, 1248–1349	*
E1A	28·5 kDa protein	576–1155, 1248–1453	*
E1A	24·7 kDa protein	576–1062, 1248–1453	*
E1A	PolyA signal for E1A	1492–1497	*
E1B	TATA box for E1B	1547–1552	*
E1B	21 kDa protein	1601–2137	*
E1B	55 kDa protein	1906–3384	*
E1B	PolyA signal for E1B	3402–3407	*
IX	TATA box for IX	3382–3387	*
IX	Hexon-associated protein IX	3478–3894	*
IX	PolyA signal for IX	3907–3912	*
IVa2	Maturation protein IVa2	(3946–5279, 5558–5570)c	*
E2B	PolyA signal for E2B	(3925–3930)c	*
E2B	TATA box for MLP	5870–5876	*
E2B	DNA polymerase	(5049–8540, 13847–13855)c	*
E2B	Terminal protein precursor (pTP)	(8420–10387, 13847–13855)c	*
VA RNA	VA RNA I	10419–10588	*
VA RNA	VA RNA II	10671–10839	*
L1	52/55 kDa protein	10868–12025	*
L1	Protein IIIa precursor	12050–13816	*
L1	PolyA signal for L1	13829–13834	*
L2	Penton base protein III	13904–15538	*
L2	Protein VII precursor	15552–16130	15550–16128
L2	L2 minor core protein pV	16173–17225	16171–17223
L2	pX 8·3 kDa protein	17254–17481	−
L2	pX 7·6 kDa protein	−	17252–17461
L2	PolyA signal for L2	17500–17505	17498–17503
L3	Protein VI precursor	17557–18309	17555–18307
L3	Hexon (protein II)	18422–21256	18420–21254
L3	23·7 kDa protease	21293–21922	21291–21920
L3	PolyA signal for L3	21942–21947	21940–21945
E2A	PolyA signal for E2A	(21954–21959)c	(21952–21957)c
E2A	DNA-binding protein	(22009–23559)c	(22057–23557)c
L4	100 kDa protein	23590–26064	23588–26062
L4	22 kDa protein	25766–26365	25764–26363
L4	33 kDa protein	25766–26234, 26284–26639	25764–26232, 26282–26637
L4	pVIII	26709–27392	26707–27390
E3	12·1 kDa protein	27392–27712	27390–27710
E3	16·1 kDa protein	27666–28106	27664–28104
E3	19·3 kDa protein	28091–28609	28089–28607
E3	20·1 kDa protein	28639–29178	28637–29176
E3	20·5 kDa protein	29191–29760	29189–29758
E3	10·3 kDa protein	29994–30269	29990–30265
E3	16·5 kDa protein	30241–30678	30237–30674

Table 1. cont.

Region	Product	Genome location	
		Strain GZ1	**Strain GZ2**
E3	15·3 kDa protein	30671–31081	30667–31077
E3	PolyA signal for E3	31114–31119	31110–31115
CDS	U exon	(31125–31286)c	(31121–31282)c
L5	Fiber protein	31301–32260	31297–32256
L5	PolyA signal for L5	32268–32273	32264–32269
E4	PolyA signal for E4	(32285–32290)c	(32281–32286)c
E4	E4 ORF6/7	(32301–32552, 33275–33448)c	(32297–32548, 33271–33444)c
E4	E4 ORF6	(32549–33448)c	(32545–33444)c
E4	E4 ORF4	(33351–33719)c	(33347–33715)c
E4	E4 ORF3	(33728–34081)c	(33724–34077)c
E4	E4 ORF2	(34078–34512)c	(34074–34508)c
E4	E4 ORF1	(34509–34886)c	(34505–34882)c
E4	TATA box for E4	(34967–34972)c	(34963–34968)c
	ITR	35163–35273	35159–35269

accession nos AY594255 and AC_0 00018) prototype strains. Like other members of the genus *Mastadenovirus*, the HAdV-3 genome is organized into early, intermediate and late transcription regions (Fig. 1). The strain GZ1 genome was 35 273 bp in length and had an overall base composition of 25·36 % A, 25·69 % C, 25·31 % G and 23·64 % T. The G + C content (51·0 mol%) was within the 50–52 mol% range noted in the literature for HAdV-B (Jin, 2001). Strain GZ1 DNA had an M_r of $2·1 \times 10^7$, determined from its base composition. The strain GZ2 genome was 35 269 bp in length and had nearly the same composition as strain GZ1. Thirty-nine protein-coding sequences and two RNA-coding sequences were identified in the genome sequences of both strains, including the pX protein (with predicted molecular masses of 8·3 kDa in strain GZ1 and 7·6 kDa in strain GZ2). Functionally, other non-coding features, such as promoters and transcription factor-binding and -recognition sites, were conserved between the two strains, as shown in Table 1.

Comparison and analysis of inverted repetitive sequences

Analysis of the two HAdV-3 strains revealed several inverted repetitive sequences (Table 2). The perfect 111 bp inverted terminal repeats (ITRs) at either end of the genome were identified in both strains. The entire HAdV-3 ITR sequences were very similar to the 108 bp ITR of the closely related HAdV-7 (strain Gomen) (Purkayastha et al., 2005) and the 114 bp ITR of HAdV-50 (strain Wan) (GenBank/EMBL accession no. AY737798), apart from seven mismatches in both cases. This contrasts with ITRs of the other HAdV-Bs: 136 bp for HAdV-3 (Tolun et al., 1979) and HAdV-7 (strain Greiner) (Shinagawa & Padmanabhan, 1980); 137 bp for HAdV-11 (Mei et al., 2003; Stone et al., 2003) and HAdV-35 (Kovács et al., 2004); and 121 bp for SAdV-21 (Davison et al., 2003). The ITRs did not contain the consensus motif CATCATCAAT found in most other HAdVs (Stone

et al., 2003). Instead, they ended with the sequence CTATCTATAT, as found in HAdV-7. The conserved TATAATATACC motif that binds the complex of terminal

Table 2. Inverted repeats in HAdV-3 strains GZ1 and GZ2

Inverted repeats in HAdV-3 strains GZ1 and GZ2 of 13 nt or longer are shown. An asterisk indicates that the same position is found in strain GZ2 as in strain GZ1. − indicates that the strain does not include the corresponding repeat.

Length (bp)	Sequence	Position	
		Strain GZ1	**Strain GZ2**
111	CTATC…CGGGG	1–111	*
	GATAG…GCCCC	35273–35163	35269–35159
14	CTGAAACTGTTTGG	−	2279–2292
	GACTTTGACAAACC	−	32436–32423
13	AAAGCGAAAGTAA	7255–7267	*
	TTTCGCTTTCATT	28130–28118	28128–28116
13	CAGCAACTTCATG	21025–21037	21023–21035
	GTCGTTGAAGTAC	21757–21745	21755–21743
13	CCTCAATCTCTTC	9196–9208	*
	GGAGTTAGAGAAG	23773–23761	23771–23759
13	CTGGTAGCCAATG	10133–10145	*
	GACCATCGGTTAC	20734–20722	20732–20720
13	GAGTTTTGGCTGG	19437–19449	19435–19447
	CTCAAAACCGACC	32837–32825	32833–32821
13	GCTGCAGCTGCTG	14927–14939	*
	CGACGTCGACGAC	21291–21279	21289–21277
13	GGAGGCAAGTCCA	8129–8141	*
	CCTCCGTTCAGGT	18122–18110	18120–18108
13	TTTCTTCTCCTTC	9622–9634	−
	AAAGAAGAGGAAG	18962–18950	−
13	TCGGGGTGAAATT	−	4252–4264
	AGCCCCACTTTAA	−	25554–25542

protein precursor (pTP) and DNA polymerase during viral DNA replication was present at 8–18 bp (Temperley & Hay, 1992). The ITRs are critical to virus replication, as well as for gene activation and transcription. The transcription factor DNA-binding motifs, such as the NFIII/Oct-1 binding site at 40–50 bp, Sp1 binding site at 72–79 bp and NFI binding site at 26–39 bp, were also identified in the ITR regions of strains GZ1 and GZ2. Other inverted repetitive sequences were short and their function is not yet known. Inverted repetitive sequences are rife in HAdV-3 genomes and perhaps they act in the transcription stages. Moreover, inverted repeats were also identified in the genomes of most other HAdV-B (data not shown).

Whole-genome comparison

Genome sequences of the two strains were aligned by using PipMaker. Both genomes shared close identity with respect to nucleotide sequences. PipMaker analysis suggested a small duplication (covering the region of nt 28 600–29 800) that is conserved in both strains and minimal differences at the gross level (Fig. 2). However, under detailed scrutiny, strain GZ2 had 93 mismatches and four gaps compared with strain GZ1, which caused typical CPE. The effect of mismatches is shown in Table 3. The substitutions in the strain GZ2 genome are obvious. Synonymous and non-synonymous substitutions resulting from mismatches are shown in Table 3. Eighteen proteins had synonymous substitutions and 27 proteins had non-synonymous substitutions, including a 'C' to 'T', which shortened the L2 pX protein (8·3 kDa in strain GZ1) to 7·6 kDa in GZ2 as a result of an internal stop codon (TAG) at nt 17459–17461. Of the 27 proteins with non-synonymous substitutions, high numbers were found in pTP (six substitutions) and

Fig. 2. Whole-genome analyses of the HAdV-3 strain GZ1 and GZ2 sequences. The genome sequences of the two strains were aligned and analysed. Dot-plot analysis of the aligned sequences was displayed by PipMaker; genome duplication (covering the region nt 28 600–29 800) is indicated by short parallel diagonal lines above and below the main line in the upper right portion of the plot.

hexon (five substitutions). Penton base, DNA Pol, 100 kDa and 12·1 kDa proteins also had three or four non-synonymous substitutions. The effects of such non-synonymous substitutions are not known. The synonymous substitutions may be single-nucleotide polymorphisms, as they did not change amino acids.

E1A and E1B

The E1A proteins regulate viral and host gene expression by interacting with various members of the host-cell transcription machinery. E1A coding sequences are conserved across the various *Mastadenovirus* species. One substitution of Y to F was identified in the E1A 28·5 kDa protein of the GZ2 genome. The E1B 55 kDa protein was similar to the large T antigen protein, which has been shown to inhibit cellular p53-mediated host-defence mechanisms (Yew et al., 1994).

E2

Three proteins required for viral DNA replication have been identified in the E2 transcriptional unit. Six non-synonymous substitutions were found in the E2B terminal protein precursor of strain GZ2, three in DNA polymerase protein and two in DNA-binding protein. The high number of substitutions in these proteins may potentially affect the replication course in the GZ2 genome.

E3

The E3 region of HAdVs encodes proteins that are not essential for *in vitro* growth. Both the 16·5 kDa and 15·3 kDa proteins were similar to adenoviral E3 proteins that are known to protect virus-infected cells against TNF-induced cytolysis (Horton et al., 1990). Non-synonymous substitutions occurred in half of the eight proteins in the E3 region, including the 15·3 kDa protein ('R to Q' and 'R to C').
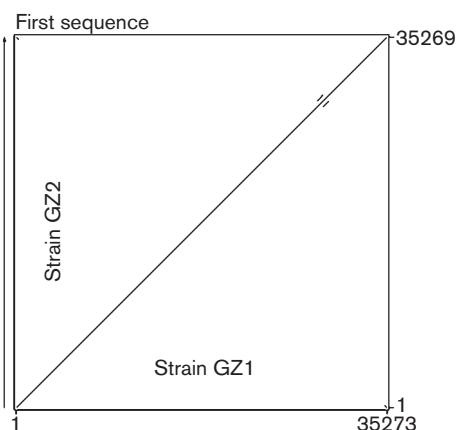
E4

Unlike the other early transcripts, the proteins encoded by the E4 transcription unit have various functions, including viral RNA export and stabilization (Leppard, 1997). Six proteins were identified in both strains. Non-synonymous substitutions were found in three proteins of both strains.

IX and IVa2

Protein IVa2 in the intermediate gene region of HAdV-3 strain GZ2 had one non-synonymous substitution. Proteins IX and IVa2 play a critical role in controlling DNA packaging during AdV assembly (Zhang et al., 2001; Sargent et al., 2004) and act as transcriptional activators depending on the presence of the TATA box upstream for HAdV-3.

L1

Protein IIIa precursor and the 52/55K protein homologue (with a predicted molecular mass of 43·8 kDa) were

Table 3. Substitutions caused by mismatches in HAdV-3 strain GZ2 coding sequences compared with strain GZ1

The absence of any substitution (synonymous or non-synonymous) in a protein is indicated by −.

Region	Product	Substitution	
		Synonymous	**Non-synonymous**
E1A	28·5 kDa protein	−	Y→F
E1B	21 kDa protein	−	Q→R
IVa2	IVa2	Y→Y	R→L
E2B	DNA Pol	Y→Y	K→E, A→V, A→V
E2B	pTP	P→P	S→F, E→G, I→T, R→G, E→G, G→G
L1	Protein IIIa precursor	Q→Q	V→L, D→N, A→V
L2	Penton base protein	V→V	I→V, D→G, M→I, G→S
L2	Minor core protein pV	−	P→L
L2	pX 7·6 kDa protein	−	Q-stop codon
L3	Protein VI precursor	−	R→K
L3	Hexon	R→R, V→V	Q→R, A→T, M→V, Y→H, T→M
L3	23·7 kDa protein	T→T	R→C
E2A	DNA-binding protein	K→K	E→G, T→P
L4	100 kDa protein	V→V, I→I, K→K	E→K, G→S, K→E
L4	33 kDa protein	H→H	Q→R, H→Y
L4	22 kDa protein	H→H	Q→R
L4	pVIII	L→L, R→R	S→N, Y→C
E3	12·1 kDa protein	−	A→S, G→E, V→D
E3	16·1 kDa protein	P→P, V→V	−
E3	19·3 kDa protein	−	G→V, G→S
E3	20·5 kDa protein	−	R→K
E3	10·3 kDa	−	R→C
E3	15·3 kDa	S→S	R→Q, R→C
CDS	U exon	V→V, G→G	G→E, T→A
L5	Fiber	−	S→P, I→M
E4	ORF6/7	L→L, I→I	A→E
E4	ORF6	D→D	C→Y
E4	ORF2	−	L→F
E4	ORF1	P→P, R→R	−

identified in both strains. Three non-synonymous substitutions were found in protein pIIIa of strain GZ2. The 52/55 kDa protein acts as a scaffold for the capsid during virus assembly (Hasson et al., 1989).

L2

Four coding sequences were identified in the L2 regions, including the penton base protein III coding sequence. Penton base protein contains a conserved Arg–Gly–Asp (RGD) sequence and is involved in virus internalization through interaction with different host integrins (Wickham et al., 1993). The penton base proteins of strains GZ1 and GZ2 were 99·7 % identical at the nucleotide level and 99·5 % at the amino acid level. There were four non-synonymous substitutions in the penton base protein of strain GZ2. The effects of these mutations are difficult to predict, as the structural and functional domains of the penton base protein have yet to be determined. A non-synonymous substitution in the pX protein, which has a predicted

molecular mass of 8·3 kDa in strain GZ1 and 7·6 kDa in strain GZ2, gave rise to an internal stop codon (TAG) at nt 17459–17461; the effects of this substitution are not yet known.

L3

Three coding sequences were found in the L3 regions: minor capsid protein precursor pVI, hexon and 23·7 kDa protease. The hexon protein accounted for 83 % of the adenovirus capsid and is known to be the principal antigenic component that results in protective immunity following natural infections. Leucine, asparagine and threonine are the three most abundant amino acids in the hexon of all HAdV-B (data not shown). A CLUSTAL-based multiple sequence alignment revealed seven HVRs (Fig. 3) between the hexons of HAdV-3, -5, -7, -11, -16, -21, -34 and -35, and SAdV-21, which account for 99 % of the serotype-specific variations (Crawford-Miksza & Schnurr, 1996). Most of the antibodies against the hexon in an adenovirus infection are directed

```
HAdV-7    121 YNSLAPKGAP NTSQWIVT-- ---------- AGEE------ RAVTTTTNTF GIASMKGDN-
SAdV-21   121 .......... ...LDKGV   TTTDNNTENG DE.DEVAEEG EEEKQA.Y.. .N.PV.AEA-
HAdV-35   121 .......... .A....AKGV PTAAAAGNGE EEH.------ TEEK.A.Y.. AN.PV.AEA-
HAdV-34   121 .......... .A...LDKGV TSTG-LVDDG NDDD-----G EEAKKA.Y.. .N.PV.AEA-
HAdV-16   121 .......... ..C..KDS-- ---------- ---------- ---DSKMH.. .V.A.P.VTG
HAdV-21   121 .......... ......AEGV KKEDG---G  SD.------  EEKNL..Y.. .N.PV.AEGG
HAdV-11   121 .......... ......AEGV KNTTG---EE HVT.------ EETN...Y.. .N.PV.AEA-
HAdV-3-1  121 .......... ........-- ---------- TNRD------ N......... .........-
HAdV-3-2  121 .......... ........-- ---------- TNRD------ N......... .........-
HAdV-5    121 ..A....... .PCE.DEAAT ALEINLEEED DDN.D-EVDE Q.EQQK.HV. .Q.PYS.I.-
                                                              HVR1

HAdV-7    161 -ITK-EGLEI GKDITADN-- -KPIYADKTY QPEPQVGEES WTDTDGTNEK FGGRALKPAT
SAdV-21   180 E...--..P. .LKVPSEG-D P.......L. .......... .....D..   Y.......E.
HAdV-35   174 K...-..P.  .LE.S.E-E  S.......L. .......D.T ..L..KT.E  Y.....T.
HAdV-34   174 K...-D..PV .LEVSTEG-- P.......L. .......D.T ..L..KT.E  Y...V...E.
HAdV-16   156 KKIEAD..P. .I.S.SGT-- DTV.....F  ....NA.    .V.AN..E.  Y....DT.
HAdV-21   171 D...DK..P. .SE..DG-E  A.......L. .......D.T ......T.   Y....D.
HAdV-11   171 E...-..PV  .LEVSDE-E  S.......L. .......L.D.T ..L..KT.  Y....D.
HAdV-3-1  161 -...-..Q.  .....TTEGE E.......... .......... .......... ..........
HAdV-3-2  161 -...-..Q.  .....TTEGE E.......... .......... .......... ..........
HAdV-5    178 -...-..IQ. .VEGQTPK-- ----...F   ....I..SQ  .YE.EIN--H AA..V..KT.
              HVR2                              HVR3

HAdV-7    217 KMKPCYGSFA RPTNIKGGQA KNRRKVKP-TE GDVETEEPDI DMEFFDGREA AD----AFSP
SAdV-21   238 .......... K...V.....  .---..KVE.  ---GKV.Y.. ..N...L.SQ KT---GLK.
HAdV-35   232 N.......Y. K...L.....  .---P.NSEP  S-S..KI.Y. .....NSSQ  RT---N...
HAdV-34   231 .......... K.........  .---..KED   DGTNNI.Y.. ..N...L.SQ RS---BLK.
HAdV-16   214 .......... K...KE....  N----LKDS.  TAAT.PNY.. ..LA..NKNI .A----NYD.
HAdV-21   229 .......... K..V.....   .---Q.TTEQ  PQNQQQV.Y. ..N..KASQ  KA----N...
HAdV-11   228 .......... K...V.....  .---Q.TTEQ  P-NQKV.Y.. .....AASQ  KT---NL..
HAdV-3-1  220 N......... ..........  .......T..  .G........ .......D   VAG---.LA.
HAdV-3-2  220 NV........ ..........  .......T..  .G........ .......D   VAG---.LA.
HAdV-5    229 P.......Y. K...EN...G  I------LVK  QQNGKL.SQV E.Q..STT.  TAGNGDNLT.
              HVR4                            HVR5

HAdV-7    272 EIVLYTENVN LETPDSHVVY KPGTSDDNSH ANLGQQAMPN RPNYIGFRDN FVGLMYYNST
SAdV-21   288 K..M.A...D .....T.....  ...A..AS.. .......S... .......... ..I......
HAdV-35   284 K..M.A...G .....T.....  ...E.TS.E  .......S... .......... ..I......
HAdV-34   284 K..M.A...D ...S.T.....  ...V..AS.E T.....S... .......... ..I......
HAdV-16   266 D..M.....D .Q...T.I..   ...E.TS.E  S......... .......... ..I......
HAdV-21   282 K..M.A...D ....T.....   ...EES.    .......S... .......... ..I......
HAdV-11   280 K..M.A.... .....T.....  ...E.TS.E  .......S... .......... ..I......
HAdV-3-1  277 .......... ..........   ...-...... .......... .......... ..........
HAdV-3-2  277 .......... ..........   ...-...... .......... .......... ..........
HAdV-5    283 KV...S.D.D I....T.IS.   M.TIKEG..R ELM...S... ......A... ..I......
              HVR6

HAdV-7    332 GNMGVLAGQA SQLNAVVDLQ DRNTELSYQL LLDSLGDRTR YFSMWNQAVD SYDPDVRIIE
SAdV-21   348 .......... .......... .......... .......... .......... .......V..
HAdV-35   344 .......... .......... .......... .......... .......... .......V..
HAdV-34   344 .......... .......... .......... .......... .......... .......V..
HAdV-16   326 .......... .......... .......... .......... .......... ..........
HAdV-21   342 .......... .......... .......... .......... .......... ..........
HAdV-11   340 .......... .......... .......... .......... .......... .......V..
HAdV-3-1  336 .......... .......... .......... .......... .......... ..........
HAdV-3-2  336 .......... .......... .......... .......... .......... ..........
HAdV-5    343 .......... .......... .......... .......... ...I...... ..........

HAdV-7    392 NHGIEDELPN YCFPLDGIG- PAKTYQGIKS K--------D NGWEKD-DNV SKSNEIAIGN
SAdV-21   408 ...V...... ........V.P RTDS.K..EN NGD------EN TT.KD--LDP NGIS.L.K.
HAdV-35   404 ...V...... ........V   .TTS.KS.VP NGE-----DN .N.KE--PR. NGTS...GQ..
HAdV-34   404 ...V...... ........V.P QTDS.KE..P NGD------Q- ST.TN--VDP NG.SQL.K..
HAdV-16   386 ...V...... .....N.V.-  FTD.....V.V TDAVAGTSG  TQ.D..DTT. .TA...HG..
HAdV-21   402 ...V...... ........V.V .ISS.KI.EP NGQ------AD .KE--PDI  NGTS...GQ..
HAdV-11   400 ...V...... ........V   .TTS.KS.VP NGD------NA PN.KE--PR. NGTS...GQ..
HAdV-3-1  396 .......... ........-  .GHR.....V .TD-----DA ......-A.  .DTA.....
HAdV-3-2  396 .......... ........-  .GHR.....V .TD-----DA ......-A.  .DTA.....
HAdV-5    403 ...T...... ....G.VI-   NTE.LTKV.P .TG----QE  ......ATEF .DK...RV..
              HVR7
```

Fig. 3. Multiple sequence alignment of the hexon proteins of HAdV-3, -5, -7, -11, -16, -21, -34 and -35, and SAdV-21; HAdV-3-1 and HAdV-3-2 correspond to strains GZ1 and GZ2, respectively. CLUSTAL_W alignment of the amino acid sequences of the hexons reveals seven major hypervariable regions (HVR1–HVR7). Dots, conserved amino acids; dashes, gaps.

against epitopes within these seven HVRs. A comparison of the hexon coding sequences from the strain GZ1 and GZ2 genomes identified two synonymous and five non-synonymous substitutions (Table 3) and they were 99 % identical at the amino acid level. Interestingly, two synonymous substitutions in strain GZ2 occurred within the HVRs, whereas the five non-synonymous substitutions were found in the conserved regions of the hexon coding sequences. The seven HVRs contained >99 % of hexon serotype-specific residues. Both strains belonged to HAdV-3, so, in these regions, a complete identity between the two strains is not unexpected. On the other hand, the hexon epitopes are known to be conformational (Crawford-Miksza & Schnurr, 1996). Therefore, a change in a structural region, such as the M221V substitution in the conserved region of

the L1 loop between HVR3 and HVR4, may affect protein folding in the antigenic regions to a certain extent.

L4

Four coding sequences were identified, corresponding to the 100 kDa, 22 kDa, 33 kDa and pVIII proteins. The 100 kDa non-structural protein is involved in hexon assembly (Oosterom-Dragon & Ginsberg, 1981), selective activation of late viral protein synthesis (Hayes et al., 1990) and inhibition of granzyme B-mediated lysis (Andrade et al., 2001). Protein VIII is associated with the formation of a possible link between the hexon capsomere and core capsid components (Shenk, 2001). In the 100 kDa protein of strain GZ2, three synonymous and three non-synonymous substitutions were identified.

L5

The adenovirus fiber protrudes from the vertices of the capsid, is responsible for the virus binding to host cells and is a major determinant of tissue tropism. The fiber coding sequences of the two strains were 99 % identical at the amino acid level, with two non-synonymous substitutions. The substitutions resulted in the amino acid changes S10P and I286M, which occurred in the fiber 'tail' and 'knob', respectively. Unlike members of other HAdV groups, members of HAdV-B do not bind the CAR (Defer et al., 1990).

Conclusion

The complete genomes of HAdV-3 strains GZ1 and GZ2 have been sequenced and annotated. The difference in CPE caused by the two strains was analysed at the genome level. Based on bioinformatic analyses, non-synonymous substitutions in the E2 terminal protein precursor, DNA polymerase protein and DNA-binding protein of strain GZ2 were identified, which may potentially affect the replication course in the strain GZ2 genome. Two non-synonymous substitutions were also identified in the GZ2 E3 15·3 kDa protein, which is similar to the proteins that protect virus-infected cells against TNF-induced cytolysis. In the conserved regions of the hexon coding sequences, five non-synonymous substitutions were found. The differential CPEs induced by the two strains must be caused by the genome differences, although this has not yet been defined precisely. Both children infected with the adenovirus strains exhibited overt disease: the child infected with GZ1 had fever and bronchitis, whereas the child with GZ2 had pharyngeal conjunctivitis. Thus, although the viruses possessed different growth characteristics in vitro, they were both virulent. Finally, as it has a different tropism from CAR-interacting HAdVs, HAdV-3 has the potential to be developed as an alternative gene-transfer vector to HAdV-5.

ACKNOWLEDGEMENTS

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.

Andrade, F., Bull, H. G., Thornberry, N. A., Ketner, G. W., Casciola-Rosen, L. A. & Rosen, A. (2001). Adenovirus L4-100K assembly protein is a granzyme B substrate that potently inhibits granzyme B-mediated cell death. *Immunity* **14**, 751–761.

Benkő, M., Harrach, B. & Russell, W. C. (2000). Family *Adenoviridae*. In *Virus Taxonomy: Seventh Report of the International Committee on Taxonomy of Viruses*, pp. 227–238. Edited by M. H. V. van Regenmortel, C. M. Fauquet, D. H. L. Bishop, E. B. Carstens, M. K. Estes, S. M. Lemon, J. Maniloff, M. A. Mayo, D. J. McGeoch, C. R. Pringle & R. B. Wickner. San Diego: Academic Press.

Benkő, M., Elő, P., Ursu, K., Ahne, W., LaPatra, S. E., Thomson, D. & Harrach, B. (2002). First molecular evidence for the existence of distinct fish and snake adenoviruses. *J Virol* **76**, 10056–10059.

Berriman, M. & Rutherford, K. (2003). Viewing and annotating sequence data with Artemis. *Brief Bioinform* **4**, 124–132.

Besemer, J. & Borodovsky, M. (1999). Heuristic approach to deriving models for gene finding. *Nucleic Acids Res* **27**, 3911–3920.

Brendel, V. & Kleffe, J. (1998). Prediction of locally optimal splice sites in plant pre-mRNA with applications to gene identification in *Arabidopsis thaliana* genomic DNA. *Nucleic Acids Res* **26**, 4748–4757.

Chany, C., Lepine, P., Lelong, M., Le-Tan-Vinh Satge, P. & Virat, J. (1958). Severe and fatal pneumonia in infants and young children associated with adenovirus infections. *Am J Hyg* **67**, 367–378.

Crawford-Miksza, L. & Schnurr, D. P. (1996). Analysis of 15 adenovirus hexon proteins reveals the location and structure of seven hypervariable regions containing serotype-specific residues. *J Virol* **70**, 1836–1844.

Davison, A. J., Wright, K. M. & Harrach, B. (2000). DNA sequence of frog adenovirus. *J Gen Virol* **81**, 2431–2439.

Davison, A. J., Benkő, M. & Harrach, B. (2003). Genetic content and evolution of adenoviruses. *J Gen Virol* **84**, 2895–2908.

Defer, C., Belin, M.-T., Caillet-Boudin, M.-L. & Boulanger, P. (1990). Human adenovirus-host cell interactions: comparative study with members of subgroups B and C. *J Virol* **64**, 3661–3673.

De Jong, J. C., Wermenbol, A. G., Verweij-Uijterwaal, M. W., Slaterus, K. W., Wertheim-Van Dillen, P., Van Doornum, G. J. J., Khoo, S. H. & Hierholzer, J. C. (1999). Adenoviruses from human immunodeficiency virus-infected individuals, including two strains that represent new candidate serotypes Ad50 and Ad51 of species B1 and D, respectively. *J Clin Microbiol* **37**, 3940–3945.

Dudding, B. A., Wagner, S. C., Zeller, J. A., Gmelich, J. T., French, G. R. & Top, F. H., Jr (1972). Fatal pneumonia associated with adenovirus type 7 in three military trainees. *N Engl J Med* **286**, 1289–1292.

Erdman, D. D., Xu, W., Gerber, S. I., Gray, G. C., Schnurr, D., Kajon, A. E. & Anderson, L. J. (2002). Molecular epidemiology of adenovirus type 7 in the United States, 1966-2000. *Emerg Infect Dis* **8**, 269–277.

Farkas, S. L., Benkő, M., Elő, P., Ursu, K., Dán, Á., Ahne, W. & Harrach, B. (2002). Genomic and phylogenetic analyses of an adenovirus isolated from a corn snake (*Elaphe guttata*) imply a common origin with members of the proposed new genus *Atadenovirus*. *J Gen Virol* **83**, 2403–2410.

Faulkner, R. & Van Rooyen, C. E. (1962). Adenoviruses types 3 and 5 isolated from the cerebrospinal fluid of children. *Can Med Assoc J* **87**, 1123–1125.

Frantzidou, F., Pavlitou, A., Mataftsi, A., Dumaidi, K. & Georgiadis, N. (2005). Molecular epidemiology of adenovirus strains isolated from patients with ocular disease in the area of Thessaloniki, Greece (1998-2002). *J Med Virol* **75**, 440–446.

Gao, W., Robbins, P. D. & Gambotto, A. (2003). Human adenovirus type 35: nucleotide sequence and vector development. *Gene Ther* **10**, 1941–1949.

Garon, C. F., Berry, K. W., Hierholzer, J. C. & Rose, J. A. (1973). Mapping of base sequence heterologies between genomes from different adenovirus serotypes. *Virology* **54**, 414–426.

Golovina, G. I., Zolotaryov, F. N. & Yurlova, T. I. (1991). Sensitive analysis of genetic heterogeneity of adenovirus types 3 and 7 in the Soviet Union. *J Clin Microbiol* **29**, 2313–2321.

Green, M., Mackey, J. K., Wold, W. S. M. & Rigden, P. (1979). Thirty-one human adenovirus serotypes (Ad1–Ad31) form five groups (A–E) based upon DNA genome homologies. *Virology* **93**, 481–492.

Hasson, T. B., Soloway, P. D., Ornelles, D. A., Doerfler, W. & Shenk, T. (1989). Adenovirus L1 52- and 55-kilodalton proteins are required for assembly of virions. *J Virol* **63**, 3612–3621.

Havenga, M. J. E., Lemckert, A. A. C., Ophorst, O. J. A. E. & 13 other authors (2002). Exploiting the natural diversity in adenovirus tropism for therapy and prevention of disease. *J Virol* **76**, 4612–4620.

Hayes, B. W., Telling, G. C., Myat, M. M., Williams, J. F. & Flint, S. J. (1990). The adenovirus L4 100-kilodalton protein is necessary for efficient translation of viral late mRNA species. *J Virol* **64**, 2732–2742.

Herbert, F. A., Wilkinson, D., Burchak, E. & Morgante, O. (1977). Adenovirus type 3 pneumonia causing lung damage in childhood. *Can Med Assoc J* **116**, 274–276.

Hierholzer, J. C. (1995). Adenoviruses. In *Diagnostic Procedures for Viral, Rickettsial, and Chlamydial Infections*, 7th edn, pp. 169–188. Edited by E. H. Lennette, D. A. Lennette & E. T. Lennette. Washington, DC: American Public Health Association.

Hong, J.-Y., Lee, H.-J., Piedra, P. A., Choi, E.-H., Park, K.-H., Koh, Y.-Y. & Kim, W.-S. (2001). Lower respiratory tract infections due to adenovirus in hospitalized Korean children: epidemiology, clinical features, and prognosis. *Clin Infect Dis* **32**, 1423–1429.

Horton, T. M., Tollefson, A. E., Wold, W. S. M. & Gooding, L. R. (1990). A protein serologically and functionally related to the group C E3 14,700-kilodalton protein is found in multiple adenovirus serotypes. *J Virol* **64**, 1250–1255.

Itakura, S., Aoki, K., Sawada, H. & Shinagawa, M. (1990). Analysis with restriction endonucleases recognizing 4- or 5-base-pair sequences of human adenovirus type 3 isolated from ocular diseases in Sapporo, Japan. *J Clin Microbiol* **28**, 2365–2369.

Itoh, N., Tanaka, K., Aoki, K., Hinokuma, R., Nakagawa, H., Takeuchi, S., Uchio, E., Shiao, S. & Ohno, S. (1999). Four new genotypes of adenovirus type 3 isolated from patients with conjunctivitis in Japan. *J Med Virol* **59**, 73–77.

Jin, Q. (2001). Adenovirus. In *Medical Molecular Virology*, pp. 691–710. Edited by Q. Jin. Beijing: Science Press.

Kajon, A. E., Murtagh, P., Garcia Franco, S., Freire, M. C., Weissenbacher, M. C. & Zorzopulos, J. (1990). A new genome type of adenovirus 3 associated with severe lower acute respiratory infection in children. *J Med Virol* **30**, 73–76.

Kajon, A. E., Mistchenko, A. S., Videla, C., Hortal, M., Wadell, G. & Avendano, L. F. (1996). Molecular epidemiology of adenovirus acute

lower respiratory infections of children in the south cone of South America (1991-1994). *J Med Virol* **48**, 151–156.

Kim, Y.-J., Hong, J.-Y., Lee, H.-J., Shin, S.-H., Kim, Y.-K., Inada, T., Hashido, M. & Piedra, P. A. (2003). Genome type analysis of adenovirus types 3 and 7 isolated during successive outbreaks of lower respiratory tract infections in children. *J Clin Microbiol* **41**, 4594–4599.

Kovács, G. M., LaPatra, S. E., D'Halluin, J. C. & Benkő, M. (2003). Phylogenetic analysis of the hexon and protease genes of a fish adenovirus isolated from white sturgeon (*Acipenser transmontanus*) supports the proposal for a new adenovirus genus. *Virus Res* **98**, 27–34.

Kovács, G. M., Davison, A. J., Zakhartchouk, A. N. & Harrach, B. (2004). Analysis of the first complete genome sequence of an Old World monkey adenovirus reveals a lineage distinct from the six human adenovirus species. *J Gen Virol* **85**, 2799–2807.

Lauer, K. P., Llorente, I., Blair, E. & 8 other authors (2004). Natural variation among human adenoviruses: genome sequence and annotation of human adenovirus serotype 1. *J Gen Virol* **85**, 2615–2625.

Leppard, K. N. (1997). E4 gene function in adenovirus, adenovirus vector and adeno-associated virus infections. *J Gen Virol* **78**, 2131–2138.

Lewis, T. C., Stout, J. W., Martinez, P., Morray, B., White, L. C., Heckbert, S. R. & Redding, G. J. (2004). Prevalence of asthma and chronic respiratory symptoms among Alaska native children. *Chest* **125**, 1665–1673.

Li, Q.-G. & Wadell, G. (1988). Comparison of 17 genome types of adenovirus type 3 identified among strains recovered from six continents. *J Clin Microbiol* **26**, 1009–1015.

Li, Q.-G., Zheng, Q. J., Liu, Y. H. & Wadell, G. (1996). Molecular epidemiology of adenovirus types 3 and 7 isolated from children with pneumonia in Beijing. *J Med Virol* **49**, 170–177.

Li, L, Shimizu, H., Doan, L. T. P. & 8 other authors (2004). Characterizations of adenovirus type 41 isolates from children with acute gastroenteritis in Japan, Vietnam, and Korea. *J Clin Microbiol* **42**, 4032–4039.

Martone, W. J., Hierholzer, J. C., Keenlyside, R. A., Fraser, D. W., D'Angelo, L. J. & Winkler, W. G. (1980). An outbreak of adenovirus type 3 disease at a private recreation center swimming pool. *Am J Epidemiol* **111**, 229–237.

Mei, Y.-F., Lindman, K. & Wadell, G. (1998). Two closely related adenovirus genome types with kidney or respiratory tract tropism differ in their binding to epithelial cells of various origins. *Virology* **240**, 254–266.

Mei, Y.-F., Skog, J., Lindman, K. & Wadell, G. (2003). Comparative analysis of the genome organization of human adenovirus 11, a member of the human adenovirus species B, and the commonly used human adenovirus 5 vector, a member of species C. *J Gen Virol* **84**, 2061–2071.

Mizuta, K., Suzuki, H., Ina, Y., Yazaki, N., Sakamoto, M., Katsushima, N. & Numazaki, Y. (1994). Six-year longitudinal analysis of adenovirus type 3 genome types isolated in Yamagata, Japan. *J Med Virol* **42**, 198–202.

Myerowitz, R. L., Stalder, H., Oxman, M. N., Levin, M. J., Moore, M., Leith, J. D., Gantz, N. M. & Hierholzer, J. C. (1975). Fatal disseminated adenovirus infection in a renal transplant recipient. *Am J Med* **59**, 591–598.

Oosterom-Dragon, E. A. & Ginsberg, H. S. (1981). Characterization of two temperature-sensitive mutants of type 5 adenovirus with mutations in the 100,000-Dalton protein gene. *J Virol* **40**, 491–500.

Purkayastha, P., Su, J., Carlisle, S., Tibbetts, C. & Seto, D. (2005). Genomic and bioinformatics analysis of HAdV-7, a human adenovirus of species B1 that causes acute respiratory disease: implications for vector development in human gene therapy. *Virology* **332**, 114–129.

Roberts, R. J., O'Neill, K. E. & Yen, C. T. (1984). DNA sequences from the adenovirus 2 genome. *J Biol Chem* **259**, 13968–13975.

Roberts, R. J., Akusjärvi, G., Aleström, P., Gelinas, R. E., Gingeras, T. R., Sciaky, D. & Pettersson, U. (1986). A consensus sequence for the adenovirus-2 genome. In *Adenovirus DNA: the Viral Genome and its Expression*, pp. 1–51. Edited by W. Doerfler. Boston: Martinus Nijhoff.

Roelvink, P. W., Lizonova, A., Lee, J. G. M., Li, Y., Bergelson, J. M., Finberg, R. W., Brough, D. E., Kovesdi, I. & Wickham, T. J. (1998). The coxsackievirus-adenovirus receptor protein can function as a cellular attachment protein for adenovirus serotypes from subgroups A, C, D, E, and F. *J Virol* **72**, 7909–7915.

Rowe, W. P., Huebner, R. J., Gilmore, L. K., Parrott, R. H. & Ward, T. G. (1953). Isolation of a cytopathogenic agent from human adenoids undergoing spontaneous degeneration in tissue culture. *Proc Soc Exp Biol Med* **84**, 570–573.

Roy, S., Gao, G., Clawson, D. S., Vandenberghe, L. H., Farina, S. F. & Wilson, J. M. (2004). Complete nucleotide sequences and genome organization of four chimpanzee adenoviruses. *Virology* **324**, 361–372.

Ryan, M. A. K., Gray, G. C., Smith, B., McKeehan, J. A., Hawksworth, A. W. & Malasig, M. D. (2002). Large epidemic of respiratory illness due to adenovirus types 7 and 3 in healthy young adults. *Clin Infect Dis* **34**, 577–582.

Sargent, K. L., Meulenbroek, R. A. & Parks, R. J. (2004). Activation of adenoviral gene expression by protein IX is not required for efficient virus replication. *J Virol* **78**, 5032–5037.

Schwartz, S., Zhang, Z., Frazer, K. A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R. & Miller, W. (2000). PipMaker – a web server for aligning two genomic DNA sequences. *Genome Res* **10**, 577–586.

Shenk, T. E. (2001). *Adenoviridae*: the viruses and their replication. In *Fields' Virology*, 4th edn, pp. 2265–2300. Edited by D. M. Knipe & P. M. Howley. Philadelphia, PA: Lippincott Williams & Wilkins.

Shiao, S., Aoki, K., Isobe, K., Tsuzuki, W. L.-P., Itoh, N., Toba, K., Kobayashi, N., Noguchi, Y. & Ohno, S. (1996). Genome analysis of adenovirus type 3 isolated in Japan. *J Clin Microbiol* **34**, 413–416.

Shields, A. F., Hackmann, R. C., Fife, K. H., Corey, L. & Meyers, J. D. (1985). Adenovirus infections in patients undergoing bone-marrow transplantation. *N Engl J Med* **312**, 529–533.

Shinagawa, M. & Padmanabhan, R. (1980). Comparative sequence analysis of the inverted terminal repetitions from different adenoviruses. *Proc Natl Acad Sci U S A* **77**, 3831–3835.

Shinagawa, M., Matsuda, A., Ishiyama, T., Goto, H. & Sato, G. (1983). A rapid and simple method for preparation of adenovirus DNA from infected cells. *Microbiol Immunol* **27**, 817–822.

Similä, S., Jouppila, R., Salmi, A. & Pohjonen, R. (1970). Encephaloningitis in children associated with an adenovirus type 7 epidemic. *Acta Paediatr Scand* **59**, 310–316.

Sirena, D., Lilienfeld, B., Eisenhut, M. & 8 other authors (2004). The human membrane cofactor CD46 is a receptor for species B adenovirus serotype 3. *J Virol* **78**, 4454–4462.

Stone, D., Furthmann, A., Sandig, V. & Lieber, A. (2003). The complete nucleotide sequence, genome organization, and origin of human adenovirus type 11. *Virology* **309**, 152–165.

Temperley, S. M. & Hay, R. T. (1992). Recognition of the adenovirus type 2 origin of DNA replication by the virally encoded DNA polymerase and preterminal proteins. *EMBO J* **11**, 761–768.

Tolun, A., Aleström, P. & Pettersson, U. (1979). Sequence of inverted terminal repetitions from different adenoviruses: demonstration of conserved sequences and homology between SA7 termini and SV40 DNA. *Cell* **17**, 705–713.

Van der Veen, J. (1963). The role of adenoviruses in respiratory disease. *Am Rev Respir Dis* **88**, 167–180.

Vogels, R., Zuijdgeest, D., van Rijnsoever, R. & 20 other authors (2003). Replication-deficient human adenovirus type 35 vectors for gene transfer and vaccination: efficient human cell infection and bypass of preexisting adenovirus immunity. *J Virol* **77**, 8263–8271.

Wadell, G. (1979). Classification of human adenoviruses by SDS-polyacrylamide gel electrophoresis of structural polypeptides. *Intervirology* **11**, 47–57.

Wadell, G. (1984). Molecular epidemiology of human adenoviruses. *Curr Top Microbiol Immunol* **110**, 191–220.

Wadell, G., Hammarskjold, M. L., Winberg, G., Varsanyi, T. M. & Sundell, G. (1980). Genetic variability of adenoviruses. *Ann N Y Acad Sci* **354**, 16–42.

Wickham, T. J., Mathias, P., Cheresh, D. A. & Nemerow, G. R. (1993). Integrins $\alpha_v\beta_3$ and $\alpha_v\beta_5$ promote adenovirus internalization but not virus attachment. *Cell* **73**, 309–319.

Wold, W. S., Mackey, J. K., Rigden, P. & Green, M. (1979). Analysis of human cancer DNA's for DNA sequence of human adenovirus serotypes 3, 7, 11, 14, 16, and 21 in group B1. *Cancer Res* **39**, 3479–3484.

Yeh, R.-F., Lim, L. P. & Burge, C. B. (2001). Computational inference of homologous gene structures in the human genome. *Genome Res* **11**, 803–816.

Yew, P. R., Liu, X. & Berk, A. J. (1994). Adenovirus E1B oncoprotein tethers a transcriptional repression domain to p53. *Genes Dev* **8**, 190–202.

Zhang, W., Low, J. A., Christensen, J. B. & Imperiale, M. J. (2001). Role for the adenovirus IVa2 protein in packaging of viral DNA. *J Virol* **75**, 10446–10454.

Zhu, B., Su, X., Gong, S., Bai, P., Zhou, R. & Liu, X. (2005). Pathogenic research of human adenovirus type 3 causing acute respiratory disease in children. *J Mod Clin Med Bioeng* **11**, 134–135.